REMARKS

This application has been reviewed in light of the Office Action dated October 4, 2003. Claims 1-4, 6-10, 12-15, 17-21, and 23-27 are presented for examination. Claims 5, 11, 16, and 22 have been canceled, without prejudice or disclaimer of the subject matter. Claims 1-4, 6-10, 12-15, 17-21, and 23 have been amended to define more clearly what Applicant regards as his invention. Claims 24-27 have been added to provide Applicant with a more complete scope of protection. Claims 1, 12, and 23-27 are in independent form. Favorable reconsideration is requested.

Applicant notes with appreciation the indication that Claims 8-10 and 19-21 would be allowable if rewritten so as not to depend from a rejected claim, and with no change in scope. These claims have not been so rewritten because, for the reasons given below, the respective base claim of each is believed to be allowable.

Claims 1-7, 11-18, 22, and 23 were rejected under 35 U.S.C. § 103(a) as being unpatentable over U.S. Patent No. 5,583,762 (*Shafer*), in view of U.S. Patent No. 6,330,574 (*Murashita*).

As shown above, Applicant has amended independent Claims 1, 12, and 23 in terms that more clearly define what he regards as his invention. Applicant submits that these amended independent claims and new independent Claims 24-27, together with the remaining claims dependent thereon, are patentably distinct from the cited prior art for at least the following reasons.

The present invention is directed to generating a document type definition of a structured document. When a document is prepared by a plurality of users, and if the individual users use tags having arbitrary titles, there is the possibility of attaching different

12

tag names to the same document element, or conversely attaching the same tag name to different elements. Thus, the semantic information attached to the tag cannot be correctly handled, and redundancy is generated with respect to the tag.

The aspects of the present invention respectively set out in the independent claims address the foregoing problem by generating a document type definition that correctly treats semantic information given to the tags.

The aspect of the present invention set forth in Claim 1 is a document type definition generating method for generating a document type definition of a structured document. The structure document contains document elements of a plurality of document element types, where each one of the plurality of document element types has a document element name and each document element has a start tag and an end tag. The method includes a physical structure judging step, a semantic structure judging step, and a document type definition generating step. The physical structure judging step judges a physical similarity between the document elements in the structured document, where the judging of the physical similarity is based on the physical position of the start tag of each document element in the structured document. The semantic structure judging step judges a semantic similarity between the document elements by comparing a character string form located between the start tag and the end tag of each of the document elements. The document type definition generating step judges a similarity of the document element tags based on the results obtained in the physical structure judging step and the semantic structure judging step, and generates the document type definition unifying the document element names of similar document elements.

Among the notable features of Claim 1 are a physical structure judging step of judging a physical similarity between the document elements in the structured document,

13

where the judging of the physical similarity is based on the physical position of the start tag of each document element in the structured document, and a semantic structure judging step of judging a semantic similarity between the document elements by comparing a character string form located between the start tag and the end tag of each of the document elements.

Shafer relates to a method for generating a grammar for a collection of sample document records and to a process for reducing the number of rules of such grammars. *Shafer* merely matches a "start tag" with the corresponding "end tag", as described on column 3, lines 14-29.

The Examiner appears to equate a physical structure judging step of judging a physical similarity between the document elements in the structured document, where performance of the judging of the physical similarity is based on each document element start tag physical position in the structured document, of Claim 1 with a *Shafer's* determining whether each acquired grammar element is combined with a rule of a given form of a hierarchical tree structure and a reduced grammar or DTD (document type definition) evolved essentially as an automatic process from the originally produced sample document grouping. Applicant respectfully disagrees with this understanding of *Shafer*. Claim 1 judges the similarity of document elements, such as "Section", "Title", etc., by comparing the physical positions of the respective start tags, whereas *Shafer* judges whether or not an acquired grammar element conforms to a rule of a given form of a hierarchical tree structure. Further, a "hierarchical tree structure" and " physical positions of each of the document element start tags" are not related to each other.

Applicant has found nothing in *Shafer* that would teach or suggest the feature of a physical structure judging step of judging a physical similarity between the document elements in the structured document, where the judging of the physical similarity is

14

based on each document element start tag physical position in the structured document, as recited in Claim 1. Furthermore, the Office Action correctly states that *Shafer* fails to disclose the semantic structure judging step of Claim 1.

For at least the above reasons, Applicant asserts that Claim 1 is clearly allowable over *Shafer*, taken alone.

Murashita relates to a technique for compressing and decompressing data. *Murashita* discusses examining tag elements in a structured document to distinguish the structures of the respective tags. However, nothing has been found in *Murashita* that would teach or suggest a semantic structure judging step of judging a semantic similarity between the document elements by comparing a character string form located between the start tag and the end tag of each of the document elements, as recited in Claim 1.

Accordingly, Claim 1 is believed clearly allowable over *Shafer* and *Murashita*, taken separately or in any proper combination.

Independent Claims 12 and 23 are apparatus and computer-readable storage medium claims respectively corresponding to method Claim 1, and are believed to be patentable for at least the same reasons as discussed above in connection with Claim 1. Additionally, independent Claims 24 and 26 include a similar feature of judging the similarity between document elements by comparing a character string form located between the start tag and the end tag of each document element, as discussed above in connection with Claim 1. Independent Claims 25 and 27 include a similar feature of judging the similarity between the document elements based on a physical similarity of the document elements according to the positions of each document element start tag in the structured document, as discussed above in connection with Claim 1. Accordingly, Claims 24-27 are believed to be patentable for at least the same reasons as discussed above in connection with Claim 1.

The other claims in this application are each dependent from one or another of the independent claims discussed above and are therefore believed patentable for the same reasons. Since each dependent claim is also deemed to define an additional aspect of the invention, however, the individual reconsideration of the patentability of each on its own merits is respectfully requested.

In view of the foregoing amendments and remarks, Applicant respectfully requests favorable reconsideration and early passage to issue of the present application.

Applicant's undersigned attorney may be reached in our New York office by telephone at (212) 218-2100. All correspondence should continue to be directed to our below listed address.

Respectfully submitted,

Attorney for Applicant

Registration No. _50,333_

FITZPATRICK, CELLA, HARPER & SCINTO
30 Rockefeller Plaza
New York, New York 10112-3801
Facsimile: (212) 218-2200
NYMAIN395080

16